We conclude our analysis of diversity among coho salmon populations with a series of phylograms or trees depicting genetic distances among samples. Three trees are presented, one for 49 unadjusted samples with 15 or more individuals (Fig. 6), one for the 33 samples formed after adjustment and homogeneity testing (Fig. 7), and the last for a subset of 27 samples (Fig. 8). In all of these trees, genetic distance is measured by Cavalli-Sforza and Edwards (1967) chord distance. The significance of nodes in these trees is tested by bootstrap analysis, in which genetic distance is estimated 1000 times among samples, using a random collection of markers, producing 1000 trees. A node is considered significant if it is recovered in more than half (500) of the bootstrap trees; bootstrap values greater than 500 are placed on the tree.

The tree, showing the relationships among 49 unadjusted samples (Fig. 6), though complex and noisy, shows considerable congruence of genetic diversity and geography. The samples from South of San Francisco (SSF) form a tight cluster. A significant node separates the Central California (CC) ESU from the Southern Oregon / Northern California (SO/NC) ESU. Samples from the SO/NA ESU are found in two significant clusters, with the exception of the Little River (Humboldt Co,) smolts, which cluster with the CC ESU. Scattered over and even outside of these clusters are the samples from Green Valley Creek of the Russian River watershed and from Redwood Creek in Marin County. Although several external nodes separating samples from the CC ESU are supported, few of the deeper nodes separating CC samples are supported.

The tree, showing the relationships of the 33 samples formed after adjustments for admixture and family structure and pooling of homogeneous samples within drainages and sites, suggests an even greater congruence of genetics and geography (Fig. 7). The SSF ESU and a large proportion of the CC ESU form significant clusters, though the node separating these two clusters is not supported. Two groups of SO/NC samples are evident, those from the Klamath / Trinity drainages, now including the Little River smolts, (though the clustering of only three of these is significant) and those from the Eel and Mattole Rivers at the southern end of the SO/NC ESU, a cluster which is recovered in 78% of the bootstrap trees. Green Valley and Redwood Creek samples remain obvious outliers on this tree. Removal of these outliers yields the final tree (Fig. 8), which strongly supports the recognized ESUs for coastal coho salmon. Significant clusters are found within each of the SO/NC, CC and SSF ESUs. Still, the node separating the CC and SSF clusters is not supported by the bootstrap analysis. Likewise, the separation of Klamath / Trinity samples from Eel / Mattole samples is not supported on this unrooted tree.

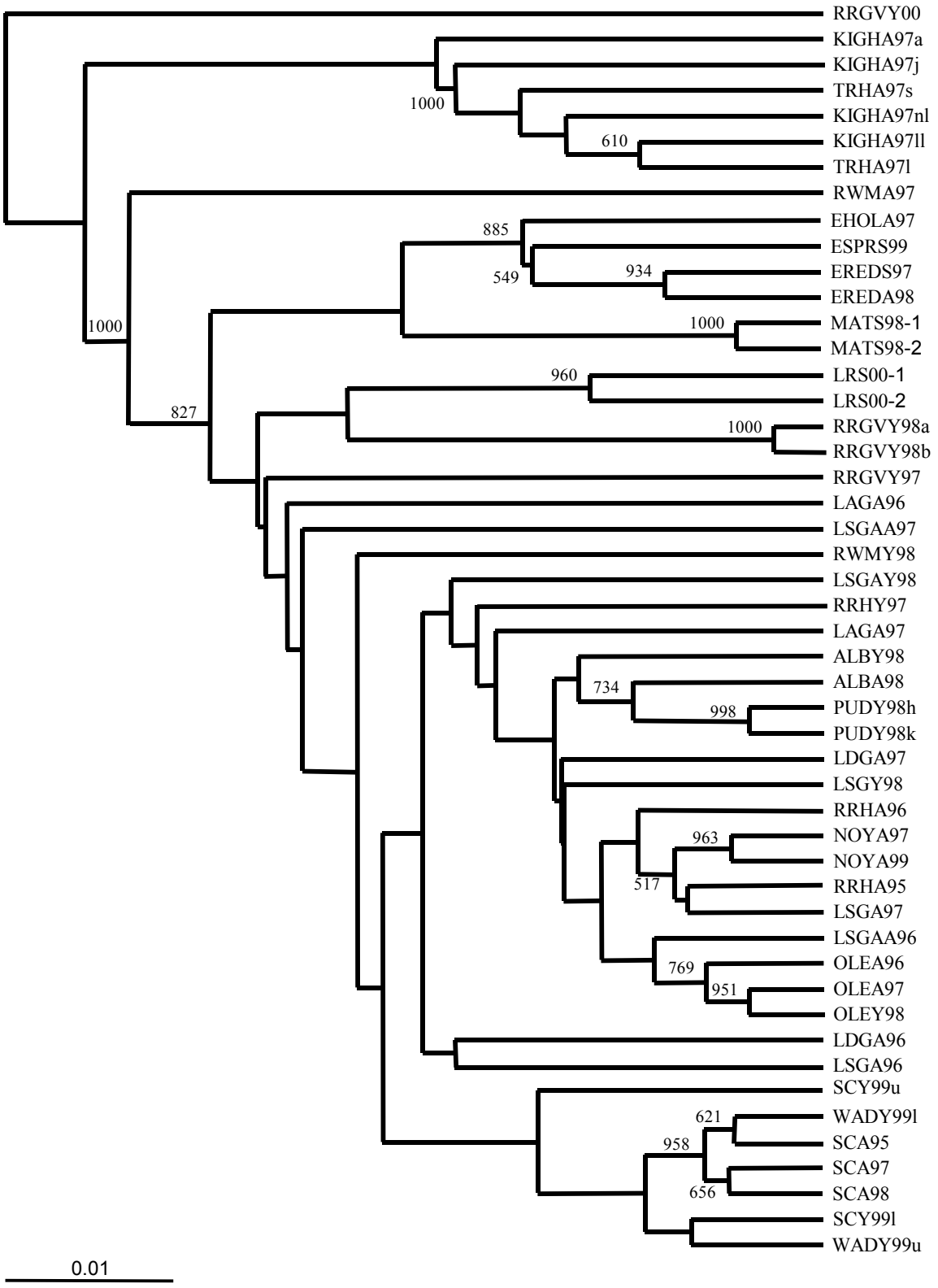

Fig. 6. Unrooted UPGMA phylogram, showing chord distances (Cavalli-Sforza and Edwards 1967) among 49 California coho salmon populations of sample size greater than 15 individuals. Nodes supported by bootstrap values greater than 500 out of 1000 are shown.

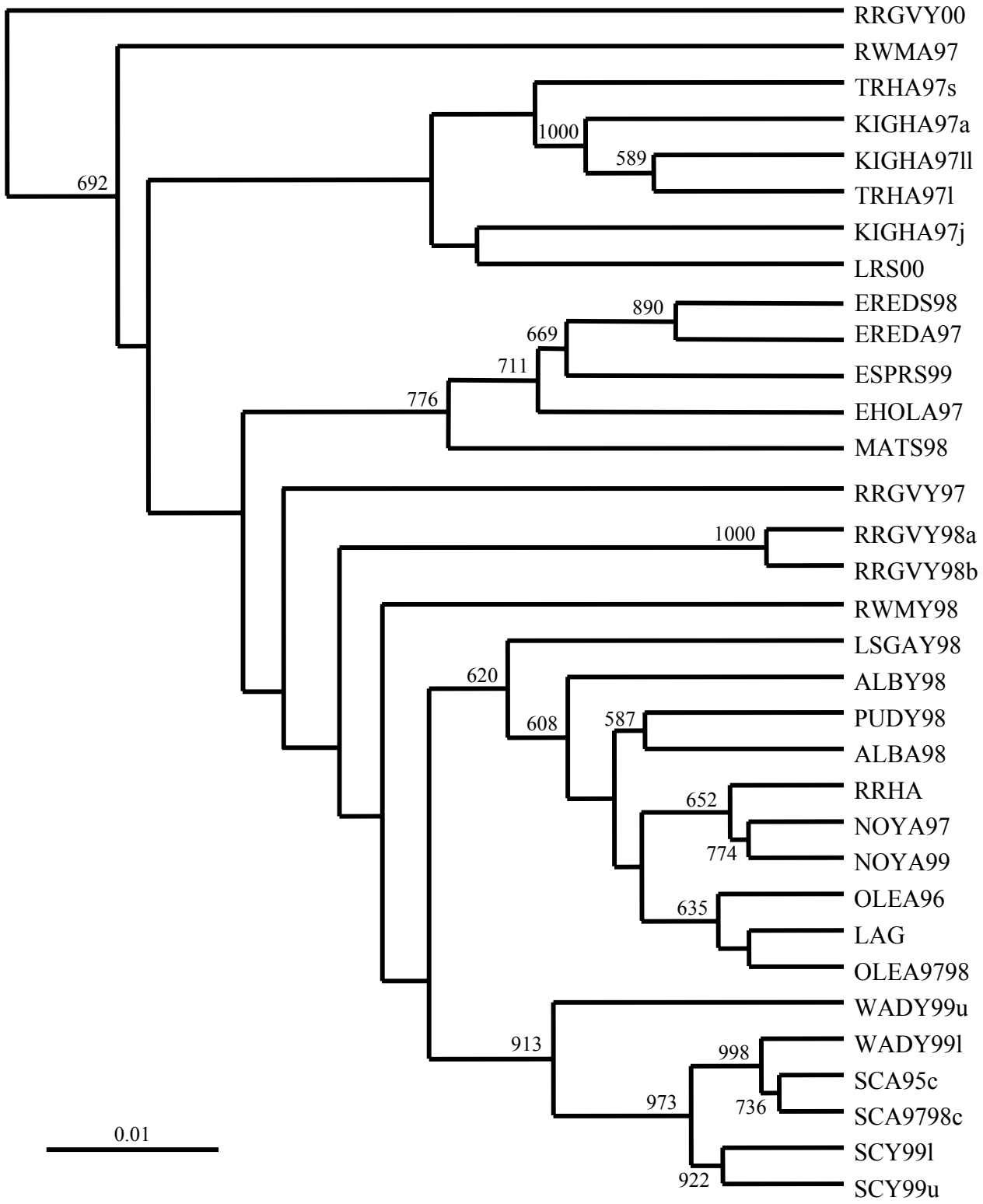

Fig. 7. An unrooted UPGMA phylogram, showing chord distances (Cavalli-Sforza and Edwards 1967) among 33 California coho salmon populations formed after adjustments for admixture and family structure and pooling of homogeneous samples within drainages and sites. Bootstrap values greater than 500 out of 1000 are shown.

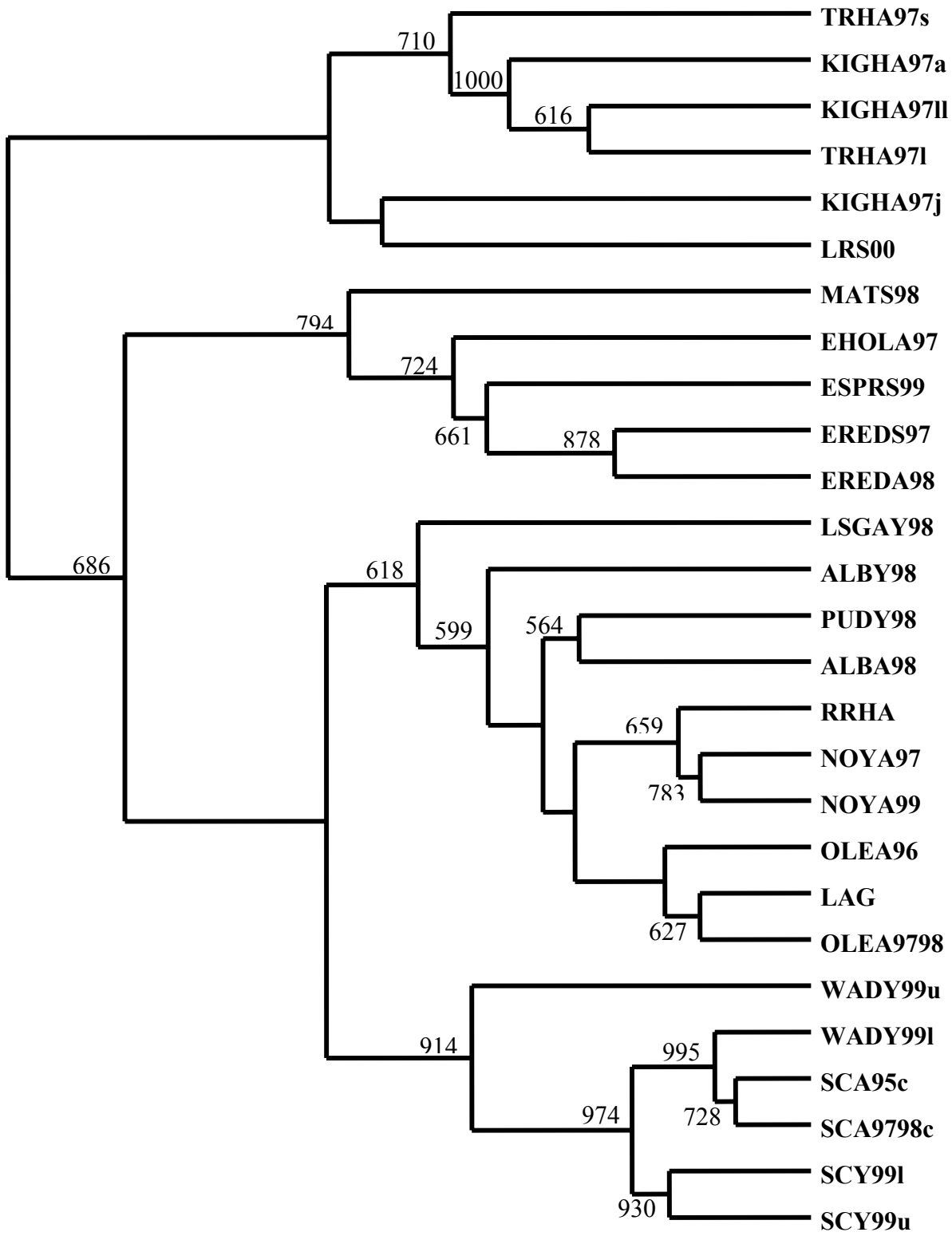

Fig. 8. An unrooted UPGMA phylogram, showing chord distances (Cavalli-Sforza and Edwards 1967) among 27 California coho salmon populations remaining after removal of Green Valley and Redwood Creek outliers on the tree in Fig. 8. Bootstrap values greater than 500 out of 1000 are shown.

Discussion

Progress towards research goals and deliverables

We contributed new knowledge relevant to all of the specific tasks in the scope for work:

1. We did determine relatedness in samples comprised of juveniles. Indeed, we went much further than that, adjusting most of these samples to correct them for family structure and to salvage them for use in describing the genetic diversity of coho salmon in Northern California.

2. We determined that temporal genetic variation among year classes is significant but smaller in magnitude than the geographical component of genetic structure.

3. We estimated significant genetic divergence among populations that was congruent with geographical distance and supportive of the present State of California ESU designations. We estimated that the effective breeding number for the Green Valley Creek population in 1998 was about 10, which raises concerns about the hatchery-based recovery program that is being based partially on captive broodstock obtained from this site.

4. We were unable to acquire historical samples to determine genetic change between historical and extant coho populations. Nevertheless, the phylogeographic structure of coho diversity suggests either that stock transfers have not erased genetic differences accumulated over evolutionary time or that the diversifying effects of genetic drift within relictual coho populations may be keeping pace with whatever homogenization has been or is being effected by hatchery practices.

5. We showed that independent environmental and biological data measured during the sampling process could be used to partition samples into subsamples that conformed better to random mating genetic equilibrium.

We elaborate on these points in the following sections.

Polymorphism of microsatellite DNA markers in coho salmon of California

We selected microsatellite DNA markers that had been developed for other species of Pacific salmon for use in the study of genetic diversity within and among coho salmon populations in California. These markers proved to be highly polymorphic, with average heterozygosities per individual ranging from 54% in a sample of juveniles from Waddell Creek to 80% in a sample of smolts from the Little River in Humboldt County. All markers are polymorphic in all populations, with the exception of *Ots-2*, which is fixed in the small sample of seven individuals collected from Green Valley Creek in 2000. The average number of alleles, which is highly dependent on sample size, ranges from 3.4 in this same Green Valley sample to 12.7 in the large pool of homogeneous samples from Lagunitas Creek. The polymorphism of the microsatellite DNA markers contrasts sharply with the low levels of protein polymorphism detected in these same coho salmon populations more than a decade ago by (Bartley et al 1992a), who reported polymorphism at only 23 of 45 loci (51%) and an average heterozygosity of only 2.7%.

The variability of these microsatellite markers makes possible the resolution of details concerning the genetics of coho salmon populations that were not possible to resolve by protein markers.

Departures from random mating equilibrium in California coho salmon populations

The distribution of genotypes within natural populations of Pacific salmon generally conform to those expected under random mating. This generalization is supported by thirty years of study of protein polymorphisms in these species (*e.g.* Bartley et al 1992b) and has been further substantiated in recent times by investigations of DNA polymorphisms (*e.g.* Banks et al 2000). Even though Bartley et al (1992a) found low variation throughout the region in protein markers, genotypic proportions at the few markers that were polymorphic did conform to those expected under random mating, and only 6.7% of the pairwise combinations of loci showed significant linkage disequilibrium.

In our study, we find widespread significant departures from random mating proportions of genotypes and more than 10% of pairwise combinations of loci showing linkage disequilibrium in nearly half of the samples formed after corrections for admixture and family structure. Part of this deviation could be attributable to residual family structure in some juvenile samples, despite our attempts to adjust for this. That family structure would be so much stronger in coho salmon populations than in samples of juvenile Chinook salmon that we were previously successful in adjusting (Banks et al 2000) suggests that the effective numbers of breeders may be quite small. Indeed, we estimate that the effective number of breeders in Green Valley Creek in 1998 may have been less than 10. Nevertheless, family structure is unlikely to explain departures from random mating genotypic proportions in adult populations, with the potential exception of small hatchery populations, such as the one in Scott Creek.

Part of the widespread deviations from random mating equilibria might be attributable to residual fine-scale Wahlund effects, *i.e.* deficiencies of heterozygotes owing to admixture in collections of individuals from populations that are genetically differentiated over small spatial scales. This seems unlikely to explain deviations in samples collected over small spatial or temporal scales, however. On the other hand, the size and significance of these departures, particularly in adult populations, suggests that these depressed populations may be experiencing inbreeding, owing to very small numbers of spawners. The finding of significant excesses of highly homozygous multi-loci genotypes in some adult populations is consistent with inbreeding. The implication of this finding is that inbreeding depression, owing to the deleterious effects of recessive lethal mutations that become homozygous upon inbreeding, just like these DNA markers have become homozygous, may be contributing to the decline in fitness of coho salmon populations in Central California.

Use of juvenile samples

In most juvenile samples, many pairs of individuals show statistically significant odds of being full brothers and sisters. Because such samples yield biased and inaccurate estimates of the genetic diversity in the adult spawning population, population geneticists in the past have avoided using juvenile samples. Nevertheless, the depressed state of coho salmon populations often precludes collections of sufficient numbers of adults. Juveniles, on the other hand, are more readily available in large numbers. Of the 57 collections available for this study, 27

comprised juveniles. To salvage these important samples for genetic analysis, we applied methods pioneered in our lab for adjusting samples for family structure to derive unbiased and accurate estimates of adult allele frequencies. Related individuals are either removed and replaced with reconstructed parents or simply removed from a sample, resulting in a sample that is smaller but usually closer to, if not in random mating equilibrium. Moreover, many these adjusted samples prove to be homogeneous with other samples from the same watershed, whereas the original sample was not. In the final phylogram used to infer the geographic distribution of genetic diversity in this study (Fig. 8), 11 of 27 populations are adjusted juvenile samples and two others are homogeneous pools that include adjusted juvenile samples. The substantial effort that juvenile samples require is repaid by the more robust inference about geographic pattern that is made possible by their use.

Temporal variation

Temporal samples or comparisons of year classes were available for seven sites: Klamath IGH, Noyo River, Russian River, Olema Creek, Lagunitas Creek, Redwood Creek (Marin Co.) and Scott Creek. Many temporal comparisons reveal significant variation. Jacks and adults were significantly different in the KIGHA samples. NOYA97 and NOYA99 were heterogeneous. The Russian River, Warm Springs Hatchery samples (RRHA95, RRHA96, RRHY97) were homogeneous but the Green Valley Creek samples were heterogeneous. The OLEA97 and OLEY98 samples were homogeneous but significantly different from OLEA96. Samples from four different years and several tributaries of Lagunitas Creek were homogeneous; only the LSGAY98 sample had to be excluded from the homogeneous LAG pool. The two samples from Redwood Creek could not be combined, even though they should represent samples from spawners (RWMA97) and offspring (RWMY98); however, these samples are outliers on the phylogram, which suggests that they are aberrant for some unknown reason. Finally, two of the Scott Creek adult samples were combinable but distinct from the third sample and from the partitioned sample of naturally spawned juveniles collected in 1999. Again, the striking deviations from random mating equilibria in these samples complicate the interpretation of temporal differences. Although temporal samples are often statistically heterogeneous, they do generally cluster closest on the phylograms, which suggests that temporal variation, though often significant is of smaller magnitude than the geographic component of genetic structure in these coho salmon populations.

Congruence of genetic diversity and geography

Bartley et al (1992a), using protein markers with low levels of polymorphism, found little congruence between genetic and geographic distances among coastal California populations of coho salmon, although they did find evidence of divergence on a larger geographic scale, between Oregon and California stocks. In our study of microsatellite DNA variation, we find genetic distances among coho samples correlating well with geographic distances among populations and strongly supporting the existing ESU designations. Given the long history of stock transfers within California and between California and other Pacific Northwest states, this congruence of genetics and geography is surprising. Two, not necessarily mutually exclusive hypotheses could explain the present spatial diversity of coho stocks in Northern California. Either the stock transfers have not "taken," owing to reduced fitness of salmon introduced via hatcheries, or the rate of population divergence has accelerated with the radical decline in the

abundance of coho salmon in the region, owing to an acceleration in genetic drift and a reduction in the absolute number of migrants between watersheds.

The implications of using Green Valley Creek coho salmon for recovery of Russian River stocks
Our finding of strong family structure in juvenile samples from Green Valley Creek has implications for the hatchery-based program aimed at recovering coho salmon populations in the Russian River watershed. Juveniles collected from Green Valley in 2001 are being reared at the Warm Springs Dam hatchery to serve as broodstock for hatchery supplementation. Because this population appears to be propagated by small numbers of breeders, perhaps as few as 10, it is quite likely that many of the juveniles collected from this creek are related to each other. Use of these fish as broodstock could accelerate inbreeding, leading to declines in population fitness and a decreasing chance of population recovery. In the 2001 annual progress report, we suggested that microsatellite genotyping could be used to help identify related broodstock and to minimize inbreeding. Our attempt to adjust for family structure based on seven microsatellite markers suggests that the reliable identification of relatives could prove very difficult unless based on a large number of DNA markers. Even if kinship could be reliably identified and inbreeding minimized, this small population appears to be anomalous and unrepresentative of the Central California ESU (see Figs. 6 and 7).

ASSESSING GENETIC VARIATION IN STEELHEAD POPULATIONS

Our scope of work listed the following objectives for steelhead: 1) to investigate the genetic consequences of migration barriers on resident populations, 2) to investigate the genetic relationship between residents and anadromous steelhead in the same watershed, 3) to investigate the genetic relationship between tributaries of the Russian River that have and have not received hatchery transplants, 4) to determine the genetic relationship of summer and winter steelhead in the Klamath and Eel rivers, which maintain large population sizes, and apply this to putative summer run stocks in the Russian River, 5) to assess whether there is evidence for widespread hatchery influence in ocean-going salmon throughout the Russian River watershed.

We began an archive of steelhead tissue samples for this research, but once the California coastal steelhead was listed federally as a threatened species, we did not have a permit to collect. Moreover, shortly after the initiation of this contract, Dr. Carlos Garza, a geneticist hired by the National Marine Fisheries Service, Santa Cruz laboratory in 2000, began a large survey of genetic variation in steelhead using microsatellite DNA markers. Rather than duplicate his effort, we focused on an alternative, though risky approach to finding markers in candidate genes for run timing differences, which was described in the 2001 annual report. This approach was discontinued after Carolyn Greig left the project for a position in Britain. The material developed by Carolyn was transferred to Dr. Michael Banks, who hopes to pursue this approach with Chinook salmon. No further effort on steelhead was made in the second year, as greater emphasis was placed on the objectives for coho and Chinook salmon.

STOCK ORIGIN ESTIMATES FOR CHINOOK JUVENILES CAPTURED IN THE RUSSIAN RIVER

This contract supported the development of baseline genetic data for Russian River Chinook salmon, permitting comparisons to Central Valley, Klamath, and Eel River stocks. The specific tasks in our scope for work were: 1) to establish a baseline of Chinook populations from Sonoma and Mendocino Counties and compare those populations to known stocks, 2) to determine the

relationship between Russian River and other coastal Chinook populations by including both extant and historical population samples from drainages such as the Eel River, 3) to continue to use and improve species identification tests developed in the first contract. Data relevant to the first two tasks is presented in this report. The species identification test, which was described in the 2001 annual progress report and by Greig et al (2002), did reveal the presence of Chinook salmon in Lagunitas Creek and did enable us to eliminate non-coho from three samples.

In a previous progress report (April 1999), we suggested that juvenile Chinook samples captured in the Russian River might not be descendants from Warm Spring Hatchery stock. We reassessed this result using seven microsatellite markers (*Ots*-2, 3, 9, 10, 104, 107 and *Oneμ* -13) and increased the sample number of both Russian River juveniles (n=78) and Warm Springs Hatchery sample sets. These results of this survey were presented in a July 2000 report, which was completed just at the beginning of this contract. Data from five river systems were analyzed: Klamath River, Trinity River, Warm Springs Hatchery (two sample sets derived from Eel River stocks), Russian River juveniles and Central Valley (winter, spring: Butte Creek, spring: Mill and Deer Creeks, fall and late fall). Genetic distance among sites show Russian River juveniles clustering with the Central Valley spring, fall and late fall populations rather than with either the two Warm Spring Hatchery populations or the Klamath/Trinity cluster. In the 2001 annual report, we cautioned that these results would have to be checked because a volunteer had initially scored the gels for the Russian River juvenile sample, and we had not yet tested and corrected for kin structure within this sample. A third problem was that the samples from the Warm Springs Hatchery showed significant departure from random mating genotypic proportions, the causes of which would need to be resolved, if possible, before their relationship to other Chinook stocks could be reliably ascertained.

Materials and Methods

We completed microsatellite analyses on 449 fish in order to assess the affinity of Russian River Chinook with other coastal Chinook populations, primarily from the Eel River (Table 9). For this report, we added 86 adults from nine mainstem Eel River samples collected by Scott Harris, CDFG. For the Russian River, we added 8 adults from Forsyth Creek and 82 smolts from Mirabel, collected by Harris and SCWA, respectively (Table 9). We compared these samples to samples of Chinook salmon from the Central Valley, which were studied by Banks et al. (2000). We also used data from samples of Chinook salmon collected in the Santa Clara Valley by the Santa Clara Valley Water District and from Chinook samples from the Klamath River, which were analyzed by Dr. Michael Banks (Oregon State University, personal communication).

DNA was extracted from samples using the Puregene™ DNA isolation kit (Gentra System), a superior extraction procedure to Chelex 100 (BioRad) particularly when extracting tissue from degraded carcasses. DNA extractions were performed using 96-well trays. We performed multiple extracts and amplifications when samples were not successfully typed.

Individuals were genotyped at up to 7 previously described unlinked microsatellite loci: *Ots-2*, *Ots-3*, *Ots-9*, and *Ots-10* (Banks et al.1999), *Ots-104* and *Ots-107* (Nelson and Beacham 1999), and *One-13* (Scribner et al. 1996). The forward PCR primer was labeled with a fluorescent phosphoamidite (HEX or fluorescein). PCR products were electrophoresed, 96 at the time with allelic controls, on a 45.0 cm wide by 22.5 cm high 8% denaturing polyacrylamide gel at 50 W

for 150 min. DNA fragments were visualized on the FMBIO® fluorescent imaging system (Hitachi Software Engineering America Ltd) and genotypes were scored with BIOIMAGE software. The data were double-checked for accuracy and independently verified by at least one other researcher. Individuals that did not produce repeatable genotypes and were difficult to score were not included in the analyses.

Table 9. List of Chinook tissue samples collected from the mainstem of the Eel river (Humboldt Co.), from the Russian River (Sonoma Co.) and from Lagunitas Creek (Marin Co.). Russian River-Warm Springs Hatchery samples originate from the Eel River[a] and the Van Arsdale[b] hatchery.

Watershed	Creek/Size Class	1997	1998	1999	2000	Total
Eel	RY/Adult	0	9*	0	0	9
	BA/Adult	0	17*	18*	0	35
	W/Adult	0	9*	0	0	9
	LV/Adult	0	9*	0	0	9
	S/Adult	0	7*	0	0	7
	T/Adult	0	0	5*	0	5
	O/Adult	0	0	6*	0	6
	BR/Adult	0	0	6*	0	6
Russian River	WS[a,b]/Adult	100[a], 94[b]	0	0	0	194
	F/Adult	0	0	8*		8
	M/Smolt	0	0	72[#]	82[#]	154
Lagunitas	Lag/Adults	0	0	0	7	7
Totals		194	51	115	89	449

Eel: (RY=Ryan; BA=Baechtal; W=Willits; LV=Long Valley; S=String; T=Tomki; O=Outlet; BR=Broaddus). Russian River: (WS=Warm Springs; F=Forsyth; M=Mirabel.
Collectors: * Harris, CDFG; [#]SCWA

We tested for deviations from Hardy-Weinberg (H-W) equilibrium within population, using GENEPOP version 3.3 (available at ftp://ftp.cefe.cnrs-mop.fr/genepop/). For linkage disequilibrium (LD), F_{IS} and F_{ST} tests we used the program GENETIX version 3.3 (available at http://www.univ-montp2.fr/~genetix/genetix.htm). The significance of F_{IS}, F_{ST} and LD ($\alpha = 0.05$) was determined by performing 500 permutations in GENETIX. We also tested genetic heterogeneity among populations from the Eel River and from the Russian River, and between Coastal populations including the Klamath River and between the Central Valley. We proceeded by measuring genetic distance between the largest homogeneous populations. The coastal populations included the Eel River, Russian River and Klamath River. The inland populations included five populations from the Central Valley (winter, spring from Butte Creek (BC), spring from Deer and Mill Creeks (DMC), fall, and late fall) and the Santa Clara Valley.

Cavalli-Sforza and Edwards (1967) (CSE) chord measures were calculated using GENDIST in the program PHYLIP (Felsenstein 1993) for data from five loci. Unweighted pair-group method